



22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1041 1042 1043 1044 1045 1046 1047 1048 1049 1050 10

atacagaaag	gcaatttttag	gaaccaaaaga	aagactgttta	agtgttttcaa	ttgtggcaaa	1980
gaaggagcaca	tagccaaaaa	ttgcagggcc	cctaggaaaa	agggctgttg	gaaatgtgga	2040
aagggaaggac	accaaatgaa	agattgtact	gagagacagg	ctaatttttt	agggaaagatc	2100
tggccttccc	acaagggaag	gccagggaat	tttcttcaga	gcagaccaga	gccaacagcc	2160
ccaccagaag	agagcttcag	gtttggggaa	gagacaacaa	ctccctctca	gaagcaggag	2220
ccgatagaca	aggaactgta	tccttttagct	tccttcagat	cactctttgg	cagcgacccc	2280
tcgtcacaa	aaagatagg	gggcaattaa	aggaagctct	attagatata	ggagcagatg	2340
atacagtatt	agaagaaatg	aatttgccag	gaagatggaa	acaaaaaatg	atagggggaa	2400
ttggaggttt	tatcaaagta	ggacagtatg	atcagatact	catagaaatc	tgcggacata	2460
aagctatagg	tacagtatta	gtaggaccta	cacctgtcaa	cataattgga	agaaatctgt	2520
tgactcagat	tggctgcact	ttaaattttc	ccattagtcc	tattgagact	gtaccagtaa	2580
aattaaagcc	aggaatggat	ggcccaaaaag	ttaaacaatg	gccattgaca	gaagaaaaaa	2640
taaaagcatt	agtagaaatt	tgtacagaaa	tggaaaagga	aggaaaaatt	tcaaaaattg	2700
ggcctgaaaa	tccatacaat	actccagtat	ttgccataaa	gaaaaaagac	agtactaaat	2760
ggagaaaatt	agtagatttc	agagaactta	ataagagAAC	tcaagatttc	tgggaagttc	2820
aattaggaat	accacatctc	gcagggttaa	aacagaaaaa	atcagtaata	gtactggatg	2880
tgggcgatgc	atattttttc	gttccttttag	ataaagactt	caggaaagta	actgcattta	2940
ccgatctgag	tataaacaat	gagaccagg	ggattagata	tcagtacaat	gtgcttcac	3000
agtagctggaa	aggatcacca	gcaatatctc	agtgtagcat	gacaaaaatc	ttagagcctt	3060
ttagaaaaaca	aaatccagac	atagtcatct	atcaatacat	ggatgatttg	tatgtaggat	3120
ctgacttaga	aatagggcag	catagaacaa	aaatagagga	actgagacaa	catctgttga	3180
ggtggggatt	taccacacca	gacaaaaaac	atcagaaaga	acctccattc	ctttggatgg	3240
gttatgaact	ccatcctgat	aaatggacag	tacagcctat	agtgtctgca	gaaaaggaca	3300
gctggactgt	caatgacata	cagaaattag	tgggaaaatt	gaattgggca	agtcagattt	3360
atgcagggat	taaagtaagg	caattatgta	aacttcttag	gggaaccaa	gcactaacag	3420
aagtagtacc	actaacagaa	gaagcagagc	tagaactggc	agaaaacagg	gagattctaa	3480
agaaccggt	acatggagtg	tattatgacc	catcaaaaga	cttaatagca	gaaatacaga	3540
agcaggggca	aggccaatgg	acatatcaaa	tttatcaaga	gccattttaa	aatctgaaaa	3600
caggaaaaata	tgcaagaatg	aagggtgcc	acactaatga	tgtgaaacaa	ttaacagagg	3660
cagtacaaaa	aatagccaca	gaaagcatag	taatatgggg	aaagactcct	aaattttaat	3720
taccataca	aaaggaaaca	tgggaagcat	ggtggacaga	gtattggcaa	gccacctgga	3780
ttcctgagtg	ggagtttgtc	aataccctc	ccttagtgaa	gttatggtac	cagttagaga	3840
aagaacccat	aataggagca	gaaactttct	atgtagatgg	ggcagccaat	agggaaacta	3900
aattaggaaa	agcaggatat	gtaactgaca	gaggaagaca	aaaagtgtgc	cccctaaccg	3960
acacaacaaa	tcagaagact	gagttacaag	caattcatct	agctttgcag	gattcgggat	4020
tagaagttaa	catagtgaca	gactcacaa	atgcattggg	aatcattcaa	gcacaaccag	4080
ataaagagtga	atcagagtta	gtcagtcaaa	taatagagca	gttaataaaa	aaggaaaaag	4140
tctacttggc	atgggtacca	gcacacaaa	gaattggagg	aaatgaacaa	gtagatgggt	4200
tggctcagtg	tggaatcagg	aaagtactat	ttttagatgg	aatagataag	gccaagaag	4260
aacatgagaa	atatcacagt	aattggagag	caatggctag	tgattttaac	ctaccacctg	4320
tagtagcaaa	agaaatagta	gccagctgtg	ataaatgtca	gctaaaaggg	gaagccatgc	4380
atggacaagt	agactgtagc	ccaggaatat	ggcagctaga	ttgtacacat	ttagaaggaa	4440
aagttatctt	ggtagcagtt	catgtagcca	gtggatatat	agaagcagaa	gtaattccag	4500
cagagacagg	gcaagaaaca	gcatacttcc	tcttaaaatt	agcaggaaga	tggccagtaa	4560
aaacagtaca	tacagacaat	ggcagcaatt	tcaccagtac	tacagttaa	gccgcctgtt	4620
ggtgggcggg	gatcaagcag	gaatttggca	ttccctacaa	tcccaaaagt	caaggagtaa	4680
tagaatctat	gaataaagaa	ttaaagaaaa	ttataggaca	ggtaagagat	caggctgaac	4740
atcttaagac	agcagtacaa	atggcagtat	tcatccacaa	ttttaaaaga	aaagggggga	4800
ttggggggta	cagtgcagg	gaaagaatag	tagacataat	agcaacagac	atacaaaacta	4860
aagaattaca	aaaacaaatt	acaaaaattc	aaaatttttcg	ggtttattac	agggacagca	4920
gagatccagt	ttggaaagga	ccagcaaagc	tcctctggaa	aggtgaagg	gcagtagtaa	4980
tacaagataa	tagtgacata	aaagtagtgc	caagaagaaa	agcaaaagatc	atcagggatt	5040
atggaaaaca	gatggcaggt	gatgattgtg	tggcaagtag	acaggtagag	gattaacaca	5100
tggaaaagat	tagtaaaaca	ccatatgtat	atttcaagga	aagctaagga	ctggttttat	5160
agacatcact	atgaaagtac	taatccaaaa	ataagttcag	aagtacacat	cccactaggg	5220
gatgctaaat	tagtaataac	aacatattgg	ggtctgcata	caggagaaa	agactggcat	5280
ttgggtcagg	gagtctccat	agaatggagg	aaaaagagat	atagcacaca	agtagaccct	5340

gacctagcag	accaactaat	tcatctgcac	tatttttgatt	gtttttcaga	atctgctata	5400
agaaatacca	tattaggacg	tatagttagt	cctaggtgtg	aatatcaagc	aggacataac	5460
aaggtaggat	ctctacagta	cttggcacta	gcagcattaa	taaaaccaaa	acagataaag	5520
ccacctttgc	ctagtgttag	gaaactgaca	gaggacagat	ggaacaagcc	ccagaagacc	5580
aagggccaca	gagggagcca	tacaatgaat	ggacactaga	gcttttagag	gaacttaaga	5640
gtgaagctgt	tagacatttt	cctaggatat	ggctccataa	cttaggacaa	catatctatg	5700
aaacttacgg	ggatacttgg	gcaggagtgg	aagccataat	aagaattctg	caacaactgc	5760
tgtttatcca	tttcagaatt	gggtgtcgac	atagcagaat	aggcgttact	cgacagagga	5820
gagcaagaaa	tggagccagt	agatcctaga	ctagagccct	ggaagcatcc	aggaagtcag	5880
cctaaaactg	cttgtagccaa	ttgtctattgt	aaaaagtggt	gcttttcattg	ccaagtttgt	5940
ttcatgacaa	aagccttagg	catctcctat	ggcaggaaga	agcggagaca	gcgacgaaga	6000
gctcatcaga	acagtcagac	tcatcaagct	tctctatcaa	agcagtaagt	agtacatgta	6060
atgcaaccta	taatagtagc	aatagtagca	ttagtagtag	caataataat	agcaatagtt	6120
gtgtggtcca	tagtaatcat	agaatatagg	aaaatattaa	gacaaagaaa	aatagacagg	6180
ttaattgata	gactaataga	aagagcagaa	gacagtggca	atgagagtga	aggagaagta	6240
tcagcacttg	tggagatggg	ggtggaaatg	gggcaccatg	ctccttggga	tattgatgat	6300
ctgtagtgt	acagaaaaat	tgtgggtcac	agtcctattat	ggggtaacctg	tgtggaagga	6360
agcaaccacc	actctatttt	gtgcctcaga	tgctaagaca	tatgatacac	aggtacataa	6420
tgtttgggcc	acacatgcct	gtgtacccac	agaccccaac	ccacaagaag	tagtattgggt	6480
aaatgtgaca	gaaaattttta	acatgtggaa	aaatgacatg	gtagaacaga	tgcattgagga	6540
tataatcagt	ttatgggatc	aaagcctaaa	gccatgtgta	aaattaaccc	cactctgtgt	6600
tagtttaaa	tgcactgatt	tgaagaatga	tactaatacc	aatagtagta	gcggggagaat	6660
gataatggag	aaaggagaga	taaaaaactg	ctctttcaat	atcagcacia	gcataagaga	6720
taaggtgcag	aaagaatatg	cattctttta	taaacttgat	atagtaccaa	tagataatac	6780
cagctatagg	ttgataagtt	gtaacacctc	agtcattaca	caggcctgtc	caaaggtatc	6840
ctttgagcca	attcccatac	attattgtgc	ccgggtggt	tttgcgattc	taaaatgtaa	6900
taataagacg	ttcaatggaa	caggaccatg	tacaaatgtc	agcacagtac	aatgtacaca	6960
tggaatcagg	ccagtagtat	caactcaact	gctgttaaat	ggcagtcctag	cagaagaaga	7020
tgtagtaatt	agatctgcca	atttcacaga	caatgctaaa	accataatag	tacagctgaa	7080
cacatctgta	gaaattaatt	gtacaagacc	caacaacaat	acaagaaaaa	gtatccgtat	7140
ccagagggga	ccaggggagag	catttgttac	aataggaaaa	ataggaaata	tgagacaagc	7200
acattgtaac	attagtagag	caaaatggaa	tgccacttta	aaacagatag	ctagcaaatt	7260
aagagaacaa	tttggaaata	ataaaacaat	aatcttttaag	caatcctcag	gaggggaccc	7320
agaaattgta	acgcacagtt	ttaattgtgg	aggggaattt	ttctactgta	attcaacaca	7380
actgtttaat	agtacttggg	ttaatatgac	ttggagtact	gaagggtcaa	ataacactga	7440
aggaagtgac	acaatcacac	tcccatgcag	aataaaacaa	ttataaaca	tgtggcagga	7500
agtaggaaaa	gcaatgtatg	ccccctccat	cagtggacaa	attagatggt	catcaaatat	7560
tactgggctg	ctattaacaa	gagatgggtg	taataacaac	aatgggtccg	agatcttcag	7620
acctggagga	ggcgatatga	gggacaattg	gagaagtga	ttatataaat	ataaagtagt	7680
aaaaattgaa	ccattaggag	tagcaccac	caaggcaaag	agaagagtgg	tgcagagaga	7740
aaaaagagca	gtgggaatag	gagctttgtt	ccttgggttc	ttgggagcag	caggaagcac	7800
tatgggctgc	acgtcaatga	cgctgacggt	acaggccaga	caattattgt	ctgatatagt	7860
gcagcagcag	aacaatttgc	tgagggtctat	tgaggcgcaa	cagcatctgt	tgcaactcac	7920
agtctggggc	atcaaacagc	tccaggcaag	aatcctggct	gtggaaagat	acctaaagga	7980
tcaacagctc	ctgggggattt	ggggttgctc	tggaaaactc	atttgcacca	ctgctgtgcc	8040
ttggaatgct	agttggagta	ataaatctct	ggaacagatt	tggaaataaca	tgacctggat	8100
ggagtgggac	agagaaaatta	acaattacac	aagcttaata	cactccttaa	ttgaagaatc	8160
gcaaaaccag	caagaaaaga	atgaacaaga	attattggaa	ttagataaat	gggcaagttt	8220
gtggaattgg	tttaacataa	caaattggt	gtggtatata	aaattattca	taatgatagt	8280
aggaggcttg	gtaggtttaa	gaatagtttt	tgtgtactt	tctatagtga	atagagttag	8340
gcagggat	tcaccattat	cgtttcagac	ccacctccca	atcccagggg	gacccgacac	8400
gcccgaagga	atagaagaag	aaggtggaga	gagagacaga	gacagatcca	ttcgattagt	8460
gaacggatcc	ttagcactta	tctgggacga	tctgctggagc	gtgtgcctct	tcagctacca	8520
cgcttgaga	gacttactct	tgattgtaac	gaggattgtg	gaacttctgg	gacgcagggg	8580
gtgggaagcg	ctcaaatatt	ggtggaatct	cctacagtat	tggagtccag	aactaaagaa	8640
tagtgctgtt	aacttgctca	atgccacagc	catagcagta	gctgagggga	cagatagggg	8700
tatagaagta	ttacaagcag	cttatagagc	tattcgccac	atacctagaa	gaataagaca	8760

450				455				460							
Ile 465	Phe	Arg	Pro	Gly	Gly 470	Gly	Asp	Met	Arg	Asp 475	Asn	Trp	Arg	Ser	Glu 480
Leu	Tyr	Lys	Tyr	Lys 485	Val	Val	Lys	Ile	Glu 490	Pro	Leu	Gly	Val	Ala 495	Pro
Thr	Lys	Ala	Lys 500	Arg	Arg	Val	Val	Gln 505	Arg	Glu	Lys	Arg	Ala 510	Val	Gly
Ile	Gly	Ala 515	Leu	Phe	Leu	Gly	Phe 520	Leu	Gly	Ala	Ala	Gly 525	Ser	Thr	Met
Gly	Cys 530	Thr	Ser	Met	Thr	Leu 535	Thr	Val	Gln	Ala	Arg 540	Gln	Leu	Leu	Ser
Asp 545	Ile	Val	Gln	Gln 550	Gln	Asn	Asn	Leu	Leu	Arg 555	Ala	Ile	Glu	Ala	Gln 560
Gln	His	Leu	Leu	Gln 565	Leu	Thr	Val	Trp	Gly 570	Ile	Lys	Gln	Leu	Gln 575	Ala
Arg	Ile	Leu	Ala 580	Val	Glu	Arg	Tyr	Leu 585	Lys	Asp	Gln	Gln	Leu 590	Leu	Gly
Ile	Trp	Gly 595	Cys	Ser	Gly	Lys	Leu 600	Ile	Cys	Thr	Thr	Ala 605	Val	Pro	Trp
Asn 610	Ala	Ser	Trp	Ser	Asn	Lys 615	Ser	Leu	Glu	Gln	Ile 620	Trp	Asn	Asn	Met
Thr 625	Trp	Met	Glu	Trp	Asp 630	Arg	Glu	Ile	Asn	Asn 635	Tyr	Thr	Ser	Leu	Ile 640
His	Ser	Leu	Ile	Glu 645	Glu	Ser	Gln	Asn	Gln 650	Gln	Glu	Lys	Asn	Glu 655	Gln
Glu	Leu	Leu	Glu 660	Leu	Asp	Lys	Trp	Ala 665	Ser	Leu	Trp	Asn	Trp 670	Phe	Asn
Ile	Thr	Asn 675	Trp	Leu	Trp	Tyr	Ile 680	Lys	Leu	Phe	Ile	Met 685	Ile	Val	Gly
Gly 690	Leu	Val	Gly	Leu	Arg	Ile 695	Val	Phe	Ala	Val	Leu 700	Ser	Ile	Val	Asn
Arg 705	Val	Arg	Gln	Gly	Tyr 710	Ser	Pro	Leu	Ser	Phe 715	Gln	Thr	His	Leu	Pro 720
Ile	Pro	Arg	Gly	Pro 725	Asp	Arg	Pro	Glu	Gly 730	Ile	Glu	Glu	Glu	Gly 735	Gly
Glu	Arg	Asp 740	Arg	Asp	Arg	Ser	Ile 745	Arg	Leu	Val	Asn	Gly 750	Ser	Leu	Ala

<212> DNA
<213> Artificial Sequence

<220>
<223> Description of the artificial sequence: synthetic DNA

<220>
<221> misc_feature
<222> (3)..(9)
<223> BstEII cleavage site

<220>
<221> misc_feature
<222> (2143)..(2148)
<223> BamHI cleavage site

<400> 9
tgggtcaccg tctattatgg ggtgcctgtg tgggaaggaag caaccaccac tctattttgt 60
gcatcagatg ctaaagcata tgatacagag gtacataatg tttgggccac acatgcctgt 120
gtacccacag accccaaccc acaagaagta gtatttgtaa atgtgacaga aaattttaac 180
atgtggaaaa atgacatggg agaacagatg catgaggata taatcagttt atgggatcaa 240
agccttaagc catgtgtaaa attaacccca ctctgtgtta gtttaaagtg cactgatttg 300
aagaatgata ctaataccaa tagtagtagc gggagaatga taatggagaa aggagagata 360
aaaaactgca gcttcaatat cagcacaagc ataagagata aggtgcagaa agaatatgca 420
ttcttttata aacttgatat agtaccaata gataatacca gctatagggt gataagttgt 480
aacacctcag tgatcacaca ggctgttcca aagggtatcct ttgagccaat tcccatacat 540
tattgtgccc cggctgggtt tgcgattcta aaatgtaata ataagacgtt caatggaaca 600
ggaccatgta caaatgtcag cacagtacaa tgtacacatg gaattcgacc agtagtatca 660
actcaactgc tgttaaattg cagtctagca gaagaagatg tagtaattag atctgccaat 720
ttcacagaca atgctaaaac cataatagta cagctgaaca catctgtaga aattaattgt 780
acaagaccca acaacaatac aagaaaaagt atccgtatcc agaggggacc agggagagca 840
tttgttacaa taggaaaaat aggaaatatg agacaagcac attgtaacat ttctagagca 900
aaatggaatg ccacttttaa acagatagct agcaaattaa gagaacaatt tggaaataat 960
aaaacaataa tctttaagca gtcacccgga ggggacccag aaattgtaac gcacagtttt 1020
aattgtggag ggggaatttt ctactgtaat tcaacacaac tgtttaatag tacttggttt 1080
aatagtactt ggagtactga aggggtcaat aacactgaag gaagtgcac aatcacactc 1140
ccatgcagaa taaaacaatt tataaacatg tggcaggaag taggaaaagc aatgtatgcc 1200
ctcccatca gtggccaaat tagatgttca tcaaataatta ctgggctgct attaaactga 1260
gatggtggta ataacaacaa tgggtccgag attttcagac ctggaggagg cgatatgagg 1320
gataattgga gaagtgaatt atataaatat aaagtagtaa aaattgaacc attaggagta 1380
gcaccacca aggcaaagag acgcgtgggt cagagagaaa agcgcgcagt gggaaatagga 1440
gctctgttcc ttgggttctt gggagcagca ggaagcacta tgggcgcagc gtcaatgacg 1500
ctgacggtac aggcagaca attattgtct gatatagtgc agcagcagaa caatttgctg 1560
agggaattg aggcgaaca gcatctgttg caactcacag tctggggcat caaacagctc 1620
caggcaagaa tcctggctgt ggaaagatac ctaaaggatc aacagctcct ggggatttgg 1680
gggtgctctg gaaaactcat ttgcaccact gctgtgcctt ggaatgctag ttggagtaat 1740
aaatctcttg aacagatttg gaataacatg acctggatgg agtgggacag agaaattaac 1800
aattacacaa gcttaataca ctcttaatt gaagaatcgc aaaaccagca agaaaagaat 1860
gaacaagaat tattggaatt agataaatgg gcaagtttgt ggaattgggt taacataaca 1920
aattggctgt ggtatataaa attattcata atgatagtag gaggcttggt aggtttaaga 1980
atagtttttg ctgtactttc tatagtgaat agagttaggc agggatattc accattatcg 2040
tttcagaccc acctccaat cccgagggga cccgacaggc ccgaaggaat agaagaagaa 2100
ggtggagaga gagacagaga cagatccatt cgattagtga acggatcc 2148

<210> 10
<211> 6229

ccacacatgc	ctgtgtaccc	acagacccca	accacaaga	agtagtattg	gtaaatgtga	1560
cagaaaat	taacatgtgg	aaaaatgaca	tggtagaaca	gatgcatgag	gatataatca	1620
gtttatggga	tcaaagccta	aagccatgtg	taaaattaac	cccactctgt	gttagtttaa	1680
agtgcactga	tttgaagaat	gatactaata	ccaatagtag	tagcgggaga	atgataatgg	1740
agaaaggaga	gataaaaaac	tgctctttca	atatcagcac	aagcataaga	gataagggtgc	1800
agaaagaata	tgcattcttt	tataaacttg	atatagtacc	aatagataat	accagctata	1860
ggttgataag	ttgtaacacc	tcagtcatta	cacaggcctg	tccaaaggta	tcctttgagc	1920
caattcccat	acattattgt	gccccggctg	gttttgcat	tctaaaatgt	aataataaga	1980
cgttcaatgg	aacaggacca	tgtacaaatg	tcagcacagt	acaatgtaca	catggaatca	2040
ggccagtagt	atcaactcaa	ctgctgttaa	atggcagctc	agcagaagaa	gatgtagtaa	2100
ttagatctgc	caatttcaca	gacaatgcta	aaaccataat	agtacagctg	aacacatctg	2160
tagaaattaa	ttgtacaaga	cccaacaaca	atacaagaaa	aagtatccgt	atccagaggg	2220
gaccagggag	agcatttgtt	acaataggaa	aaataggaaa	tatgagacaa	gcacattgta	2280
acattagtag	agcaaaatgg	aatgccactt	taaaacagat	agctagcaaa	ttaagagaac	2340
aatttgaaaa	taataaaaaca	ataatcttta	agcaatcctc	aggaggggac	ccagaaattg	2400
taacgcacag	ttttaattgt	ggaggggaat	ttttctactg	taattcaaca	caactgttta	2460
atagtacttg	gtttaatagt	acttgagta	ctgaagggtc	aaataacact	gaagggaagt	2520
acacaatcac	actcccctgc	agaataaaac	aattttataaa	catgtggcag	gaagtaggaa	2580
aagcaatgta	tgccccctcc	atcagtgga	aaattagatg	ttcatcaaat	attactgggc	2640
tgctattaac	aagagatggg	ggtaataaca	acaatgggtc	cgagatcttc	agacctggag	2700
gaggcgatat	gagggacaat	tggagaagtg	aattatataa	atataaagta	gtaaaaattg	2760
aaccattagg	agtagcacc	accaaggcaa	agagaagagt	ggtgcagaga	gaaaaaagag	2820
cagtgggaat	aggagctttg	ttccttgggt	tcttgggagc	agcaggaagc	actatgggct	2880
gcacgtcaat	gacgctgacg	gtacaggcca	gacaattatt	gtctgatata	gtgcagcagc	2940
agaacaat	gctgagggct	attgagggcg	aacagcatct	gttgcaactc	acagtctggg	3000
gcatcaaaca	gctccaggca	agaatcctgg	ctgtggaaag	atacctaaag	gatcaacagc	3060
tcctggggat	ttgggggttg	tctggaaaac	tcatttgcac	cactgctgtg	ccttggaatg	3120
ctagtgtggag	taataaatct	ctggaacaga	tttggaataa	catgacctgg	atggagtggg	3180
acagagaaat	taacaattac	acaagcttaa	tacactcctt	aattgaagaa	tcgcaaaacc	3240
agcaagaaaa	gaatgaacaa	gaattattgg	aattagataa	atgggcaagt	ttgtggaatt	3300
ggtttaacat	aacaaattgg	ctgtggtata	taaaattatt	cataatgata	gtaggaggct	3360
tggtagggtt	aagaatagtt	tttgctgtac	tttctatagt	gaatagagtt	aggcagggat	3420
attcaccatt	atcgtttcag	accacacctc	caatcccag	gggacccgac	aggccccgaag	3480
gaatagaaga	agaagggtgga	gagagagaca	gagacagatc	cattcgatta	gtgaacggat	3540
ccttagcact	tatctgggac	gatctgcgga	gcctgtgcct	cttcagctac	caccgcttga	3600
gagacttact	cttgattgta	acgaggattg	tggaaacttct	gggacgcagg	gggtgggaag	3660
ccctcaaata	ttggtggaat	ctcctacagt	attggagtca	ggaactaaag	aatagtgtctg	3720
ttaacttgct	caatgccaca	gccatagcag	tagctgaggg	gacagatagg	gttatagaag	3780
tattacaagc	agcttataga	gctattcgcc	acatacctag	agaataaaga	cagggcttgg	3840
aaaggatttt	gctataagat	gggtggcaag	tggtcaaaaa	gtagtgtgat	tggatggcct	3900
gctgtaaggg	aaagaatgag	acgagctgag	ccagcagcag	atgggggtggg	agcagtatct	3960
cgagatctag	actagaacta	gcttcgatcc	agacatgata	agatacattg	atgagtttgg	4020
acaaaccaca	actagaatgc	agtgaaaaaa	atgctttatt	tgtgaaat	gtgatgctat	4080
tgctttat	gtaaccatta	taagctgcaa	taaacaagtt	aacaacaaca	attgcattca	4140
ttttatgt	cagggttcagg	gggaggtgtg	ggaggttttt	taaagcaagt	aaaacctcta	4200
caaatgtgg	atggctgatt	atgatcctgc	ctcgcgctt	tcggtgatga	cggtgaaaac	4260
ctctgacaca	tgcagctccc	ggagacgggtc	acagcttgct	tgtaaagcga	tgccgggagc	4320
agacaagccc	gtcagggcgc	gtcagcgggt	gttggcgggt	gtcggggcgc	agccatgacc	4380
cagtcacgta	gcgatagcgg	agtgtatact	ggcttaacta	tgccggcatca	gagcagattg	4440
tactgagagt	gcaccatatg	tcgggcccgcg	ttgctggcgt	ttttccatag	gctccgcccc	4500
cctgacgagc	atcacaaaaa	tcgacgctca	agtcagaggt	ggcgaaaacc	gacaggacta	4560
taaagatacc	aggcggtttcc	ccctggaagc	tcctcgtgct	gctctcctgt	tccgacctg	4620
ccgcttaccg	gatacctgtc	cgcctttctc	ccttcgggaa	gcgtggcgct	ttctcatagc	4680
tcacgctgta	ggtatctcag	ttcggtgtag	gtcgttcgct	ccaagctggg	ctgtgtgcac	4740
gaaccccccg	ttcagcccga	ccgctgcgcc	ttatccggta	actatcgtct	tgagtccaac	4800
ccggttaagac	acgacttatc	gccactggca	gcagccactg	gtaacaggat	tagcagagcg	4860
aggatatgtag	gcggtgctac	agagttcttg	aagtggtggc	ctaactacgg	ctacactaga	4920

```

aggacagtat ttggtatctg cgctctgctg aagccagtta ccttcggaaa aagagttggt 4980
agctcttgat cccgcaaaaca aaccaccgct ggtagcgggtg gtttttttgt ttgcaagcag 5040
cagattacgc gcagaaaaaaa aggatctcaa gaagatcctt tgatcttttc tacgggggtct 5100
gacgctcagt ggaacgaaaaa ctcacgttaa gggatttttg tcatgagatt atcaaaaagg 5160
atcttcacct agatcctttt aaattaaaaa tgaagtttta aatcaatcta aagtatatat 5220
gagtaaaactt ggtctgacag ttaccaatgc ttaatcagtg aggcacctat ctcagcgatc 5280
tgtctatttc gttcatccat agttgcctga ctccccgtcg tgtagataac tacgatacgg 5340
gagggccttac catctggccc cagtgcctga atgataccgc gagaccacg ctcaccggct 5400
ccagatttat cagcaataaa ccagccagcc ggaagggccg agcgcagaag tggctcctgca 5460
actttatccg cctccatcca gtctattaat tgttgccggg aagctagagt aagtagttcg 5520
ccagttaata gtttgcgcaa cgttggtgccc attgctacag gcatcgtggt gtcacgctcg 5580
tcgtttggta tggcttcatt cagctccggt tcccaacgat caaggcgagt tacatgatcc 5640
cccattgtgt gcaaaaaagc ggtagctcc ttcggctctc cgatcgttgt cagaagtaag 5700
ttggcccgag tgttatcact catggttatg gcagcactgc ataattctct tactgtcatg 5760
ccatccgtaa gatgcttttc tgtgactggt gagtactcaa ccaagtcatt ctgagaatag 5820
tgtatgccc gaccgagttg ctcttgccc ggcgcaatac gggataatac cgcgccacat 5880
agcagaactt taaaagtgt catcattgga aaacgttctt cggggcgaaa actctcaagg 5940
atcttaccgc tgttgagatc cagttcgatg taaccactc gtgcacccaa ctgatcttca 6000
gcatctttta ctttcaccag cgtttctggg tgagcaaaaa caggaaggca aaatgcccga 6060
aaaaagggaa taagggcgac acggaaatgt tgaatactca tactcttctt ttttcaatat 6120
tattgaagca tttatcaggg ttattgtctc atgagcggat acatatttga atgtatttag 6180
aaaaataaac aaataggggt tccgcgcaca tttccccgaa aagtgccac 6229

```

```

<210> 11
<211> 860
<212> DNA
<213> Human immunodeficiency virus

<220>
<221> misc_feature
<222> (1)..(860)
<223> PI-932 original sequence V1-V2-V3-loop

```

```

<400> 11
tgtgtaccca cagaccccaa cccacaaaag gtagtattgg aaaatgtgac agaaaatttt 60
aacatgtgga aaaatgacat ggtagaacag atgcatgagg atataatcaa tttatgggat 120
caaagcctaa agccatgtgt aaaactaacc ccactctgtg ttacttttaa ttgcactgat 180
gctgatttaa attgcaataa tactgattta aattgcacta aagctaattt ggggaaaaat 240
actcataaca atactattag tgggaaaata atagagaaag tagaaataaa aaactgctct 300
ttcaagggtca ccacaggcat aagggataag atgcaaaaag aatatgcact tttgaataaa 360
cttgatatag taccaataga taatgataag aataatacta actttatatt gataagttgt 420
aacacctcga ccattacaca ggctgttcca aaggatatct ttgagccaat tcccatatcat 480
ttttgtgccc cggctggttt tgcgattcta aagtgtaatg aaaagagtta cagtggaaaa 540
ggaccatgta aaaatgtcag cacagtacaa tgtacacatg gaattaggcc agtagtgtca 600
actcaactgc tgttgaatgg cagtctagca gaaaaagaag tagtaattag atctgagaat 660
ttcacagaca atgctaaaac cataatagta cagctgaagg aatctgtaaa cattacttgt 720
ataagacccc acaacactgt aacagacagg atacatatag ggccaggagg atcatttcat 780
acaacaagaa aaataaaagg agatataaga caagcacatt gtagccttag gagaaaagat 840
tggataaaca ctttacaaga                                     860

```

```

<210> 12
<211> 870
<212> DNA
<213> Artificial Sequence

```

<220>

<223> Description of the artificial sequence: PI-932
gene cassette, comprising the cleavage sites for
restriction enzymes BspT1, PstI, BclI, EcoRI,
BglII, PvuII, XbaII, NheI

<400> 12

```
tgtgtaccca cagaccccaa cccacaaaag gtagtattgg aaaatgtgac agaaaatttt 60
aacatgtgga aaaatgacat ggtagaacag atgcatgagg atataatcaa tttatgggat 120
caaagcctta agccatgtgt aaaactaacc ccactctgtg ttactttaaa ttgcactgat 180
gctgatttaa attgcaataa tactgattta aattgcacta aagctaattt ggggaaaaat 240
actcataact gcagtattag tgggaaaata atagagaaag tagaaataaa aaactgctct 300
ttcaagggtca ccacaggcat aagggataag atgcaaaaag aatatgcact tttgaataaa 360
cttgatatag taccaataga taatgataag aataatacta actttatatt gataagttgt 420
aacacctcgg tgatcacaca ggctgtcca aaggatcct ttgagccaat tcccatacat 480
ttttgtgccc cggctgggtt tgcgattcta aagtgtaatg aaaagagtta cagtggaaaa 540
ggaccatgta aaaatgtcag cacagtacaa tgtacacatg gaattcggcc agtagtgtca 600
actcaactgc tgttgaatgg cagtctagca gaaaaagaag tagtaattag atctgagaat 660
ttcacagaca atgctaaaac cataatagta cagctgaagg aatctgtaaa cattacttgt 720
ataagacccc acaacactgt aacagacagg atacatatag ggccagggag atcatttcat 780
acaacaagaa aaataaaaagg agatataaga caagcacatt gtagcctttc tagaaaagat 840
tggaataaca ctttacaaga gatagctagc 870
```

PI-932" GATCCTGG